

OLPE  
2/10/2001 #3

Serial Number: 09/875,494

CRF Processing Date: 2/10/2001  
Edited by: [Signature]  
Verified by: [Signature] (STIC staff)

ENTERED

- ☐ Changed a file from non-ASCII to ASCII
- ☐ Changed the margins in cases where the sequence text was "wrapped" down to the next line.
- ☐ Edited a format error in the Current Application Data section, specifically: \_\_\_\_\_
- ☐ Edited the Current Application Data section with the actual current number. The number inputted by the applicant was ☐ the prior application data; or ☐ other \_\_\_\_\_
- ☐ Added the mandatory heading and subheadings for "Current Application Data".
- ☐ Edited the "Number of Sequences" field. The applicant spelled out a number instead of using an integer.
- ☐ Changed the spelling of a mandatory field (the headings or subheadings), specifically: \_\_\_\_\_
- ☐ Corrected the SEQ ID NO when obviously incorrect. The sequence numbers that were edited were: \_\_\_\_\_
- ☐ Inserted or corrected a nucleic number at the end of a nucleic line. SEQ ID NO's edited: \_\_\_\_\_
- ☐ Corrected subheading placement. All responses must be on the same line as each subheading. If the applicant placed a response below the subheading, this was moved to its appropriate place.
- ☐ Inserted colons after headings/subheadings. Headings edited included: \_\_\_\_\_
- ☐ Deleted extra, invalid, headings used by an applicant, specifically: \_\_\_\_\_
- ☒ Deleted: ☒ non-ASCII "garbage" at the beginning/end of files; ☐ secretary initials/lifename at end of file; ☐ page numbers throughout text; ☐ other invalid text, such as \_\_\_\_\_
- ☐ Inserted mandatory headings, specifically: \_\_\_\_\_
- ☐ Corrected an obvious error in the response, specifically: \_\_\_\_\_
- ☐ Edited identifiers where upper case is used but lower case is required, or vice versa.
- ☐ Corrected an error in the Number of Sequences field, specifically: \_\_\_\_\_
- ☐ A "Hard Page Break" code was inserted by the applicant. All occurrences had to be deleted.
- ☐ Deleted *ending* stop codon in amino acid sequences and adjusted the "(A)Length:" field accordingly (error due to a PatentIn bug). Sequences corrected: \_\_\_\_\_
- ☐ Other: \_\_\_\_\_

\*Examiner: The above corrections must be communicated to the applicant in the first Office Action. DO NOT send a copy of this form.

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/875,494

DATE: 07/10/2001

TIME: 11:51:32

Input Set : A:\Pto.amc

Output Set: N:\CRF3\07102001\I875494.raw

```

5 <110> APPLICANT: Schmidt, Brian
6     Chen, Tseng-hui
8 <120> TITLE OF INVENTION: Novel Leader Peptides for Enhancing Secretion of Recombinant
Protein from
9     a Host Cell
11 <130> FILE REFERENCE: COUL-012-01
C--> 13 <140> CURRENT APPLICATION NUMBER: US/09/875,494
C--> 13 <141> CURRENT FILING DATE: 2001-06-05
13 <150> PRIOR APPLICATION NUMBER: US 60/209,517
14 <151> PRIOR FILING DATE: 2000-06-05
16 <160> NUMBER OF SEQ ID NOS: 25
18 <170> SOFTWARE: PatentIn version 3.0
20 <210> SEQ ID NO: 1
21 <211> LENGTH: 24
22 <212> TYPE: PRT
C--> 23 <213> ORGANISM: Artificial
25 <220> FEATURE:
26 <223> OTHER INFORMATION: synthetic peptide
28 <400> SEQUENCE: 1
30 Met Ala Lys Lys Asn Ser Thr Leu Leu Val Ala Val Ala Ala Leu Ile
31 1             5             10             15
33 Phe Met Ala Gly Arg Ala Asn Ala
34             20
36 <210> SEQ ID NO: 2
37 <211> LENGTH: 24
38 <212> TYPE: PRT
C--> 39 <213> ORGANISM: Artificial
41 <220> FEATURE:
42 <223> OTHER INFORMATION: synthetic peptide
44 <400> SEQUENCE: 2
46 Met Ala Lys Lys Asn Ser Thr Leu Leu Val Ala Val Ala Ala Leu Ile
47 1             5             10             15
49 Met Phe Thr Gln Pro Ala Asn Ala
50             20
52 <210> SEQ ID NO: 3
53 <211> LENGTH: 24
54 <212> TYPE: PRT
C--> 55 <213> ORGANISM: Artificial
57 <220> FEATURE:
58 <223> OTHER INFORMATION: synthetic peptide
60 <400> SEQUENCE: 3
62 Met Gly Lys Lys Gln Thr Ala Val Ala Phe Ala Leu Ala Leu Leu Ala
63 1             5             10             15
65 Leu Ser Met Thr Pro Ala Tyr Ala
66             20
68 <210> SEQ ID NO: 4
69 <211> LENGTH: 24
70 <212> TYPE: PRT

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/875,494

DATE: 07/10/2001

TIME: 11:51:32

Input Set : A:\Pto.amc

Output Set: N:\CRF3\07102001\I875494.raw

```

C--> 71 <213> ORGANISM: Artificial
      73 <220> FEATURE:
      74 <223> OTHER INFORMATION: synthetic peptide
      76 <400> SEQUENCE: 4
      78 Met Gly Arg Lys Gln Thr Ala Val Ala Phe Ala Leu Ala Leu Leu Ser
      79 1          5          10          15
      81 Leu Ala Phe Thr Asn Ala Tyr Ala
      82          20
      84 <210> SEQ ID NO: 5
      85 <211> LENGTH: 106
      86 <212> TYPE: DNA
C--> 87 <213> ORGANISM: Artificial
      89 <220> FEATURE:
      90 <223> OTHER INFORMATION: synthetic
      92 <400> SEQUENCE: 5
      93 acccggttttt ttgggctaac aggaggaatt aaccatggct aaaaagaact ccaccctgct      60
      95 cgttgcagta gctgcgctga tcttcatggc cggaagggcc aacgct      106
      98 <210> SEQ ID NO: 6
      99 <211> LENGTH: 106
     100 <212> TYPE: DNA
C--> 101 <213> ORGANISM: Artificial
     103 <220> FEATURE:
     104 <223> OTHER INFORMATION: synthetic
     106 <400> SEQUENCE: 6
     107 acccggttttt ttgggctaac aggaggaatt aaccatggct aaaaagaact ccaccctgct      60
     109 cgttgcagta gctgcgctta tcatgttcac tcagccggcg aacgct      106
     112 <210> SEQ ID NO: 7
     113 <211> LENGTH: 106
     114 <212> TYPE: DNA
C--> 115 <213> ORGANISM: Artificial
     117 <220> FEATURE:
     118 <223> OTHER INFORMATION: synthetic
     120 <400> SEQUENCE: 7
     121 acccggttttt ttgggctaac aggaggaatt aaccatgggt aagaaacaga ccgctgttgc      60
     123 attcgctctg gcgctcctgg ctctttctat gaccccggcg tacgct      106
     126 <210> SEQ ID NO: 8
     127 <211> LENGTH: 106
     128 <212> TYPE: DNA
C--> 129 <213> ORGANISM: Artificial
     131 <220> FEATURE:
     132 <223> OTHER INFORMATION: synthetic
     134 <400> SEQUENCE: 8
     135 acccggttttt ttgggctaac aggaggaatt aaccatgggt cgtaaacaga ccgcagtagc      60
     137 attcgctctt gcgctgcttt ctctcgcttt caccaacgcg tacgct      106
     140 <210> SEQ ID NO: 9
     141 <211> LENGTH: 74
     142 <212> TYPE: DNA
C--> 143 <213> ORGANISM: Artificial
     145 <220> FEATURE:

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/875,494

DATE: 07/10/2001

TIME: 11:51:32

Input Set : A:\Pto.amc

Output Set: N:\CRF3\07102001\I875494.raw

```

146 <223> OTHER INFORMATION: synthetic
148 <400> SEQUENCE: 9
149 ccatgggctaa aaagaactcc accctgctcg ttgcagtagc tgcgctgac ttcatggccg      60
151 gaagggccaa cgct                                                         74
154 <210> SEQ ID NO: 10
155 <211> LENGTH: 74
156 <212> TYPE: DNA
C--> 157 <213> ORGANISM: Artificial
159 <220> FEATURE:
160 <223> OTHER INFORMATION: synthetic
162 <400> SEQUENCE: 10
163 ccatgggctaa aaagaactcc accctgctcg ttgcagtagc tgcgcttacc atgttcaactc      60
165 agccggcgaa cgct                                                         74
168 <210> SEQ ID NO: 11
169 <211> LENGTH: 74
170 <212> TYPE: DNA
C--> 171 <213> ORGANISM: Artificial
173 <220> FEATURE:
174 <223> OTHER INFORMATION: synthetic
176 <400> SEQUENCE: 11
177 ccatgggctcg taaacagacc gctgttgcac tcgctctggc gctcctgtct cttgctttca      60
179 ccaacgcgta cgct                                                         74
182 <210> SEQ ID NO: 12
183 <211> LENGTH: 74
184 <212> TYPE: DNA
C--> 185 <213> ORGANISM: Artificial
187 <220> FEATURE:
188 <223> OTHER INFORMATION: synthetic
190 <400> SEQUENCE: 12
191 ccatgggctaa gaaacagacc gctgttgcac tcgctctggc gctcctggct ctttctatga      60
193 ccccggcgta cgct                                                         74
196 <210> SEQ ID NO: 13
197 <211> LENGTH: 74
198 <212> TYPE: DNA
C--> 199 <213> ORGANISM: Artificial
201 <220> FEATURE:
202 <223> OTHER INFORMATION: synthetic
204 <400> SEQUENCE: 13
205 ccatgggctaa gaaacagacc gctgttgcac tcgctctggc gctcctgtct cttgctttca      60
207 ccccggcgta cgct                                                         74
210 <210> SEQ ID NO: 14
211 <211> LENGTH: 24
212 <212> TYPE: PRT
C--> 213 <213> ORGANISM: Artificial
215 <220> FEATURE:
216 <223> OTHER INFORMATION: synthetic peptide
218 <400> SEQUENCE: 14
220 Met Gly Lys Lys Gln Thr Ala Val Ala Phe Ala Leu Ala Leu Leu Ser
221 1          5          10          15

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/875,494

DATE: 07/10/2001

TIME: 11:51:32

Input Set : A:\Pto.amc

Output Set: N:\CRF3\07102001\I875494.raw

```

223 Leu Ala Phe Thr Pro Ala Tyr Ala
224      20
226 <210> SEQ ID NO: 15
227 <211> LENGTH: 74
228 <212> TYPE: DNA
C--> 229 <213> ORGANISM: Artificial
231 <220> FEATURE:
232 <223> OTHER INFORMATION: synthetic
234 <400> SEQUENCE: 15
235 ccatgggtcg taaacagacc gcagtagcat tcgctcttgc gctgctttct ctcgctttca      60
237 ccaacgcgta cgct      74
240 <210> SEQ ID NO: 16
241 <211> LENGTH: 69
242 <212> TYPE: DNA
243 <213> ORGANISM: Escherichia coli
245 <400> SEQUENCE: 16
246 atgaaaaaga atatcgcat tcttcttgca tctatgttcg ttttttctat tgctacaaac      60
248 gcgtacgt      69
251 <210> SEQ ID NO: 17
252 <211> LENGTH: 23
253 <212> TYPE: PRT
254 <213> ORGANISM: Escherichia coli
256 <400> SEQUENCE: 17
258 Met Lys Lys Asn Ile Ala Phe Leu Leu Ala Ser Met Phe Val Phe Ser
259 1      5      10      15
261 Ile Ala Thr Asn Ala Tyr Ala
262      20
264 <210> SEQ ID NO: 18
265 <211> LENGTH: 68
266 <212> TYPE: DNA
267 <213> ORGANISM: Erwinia carotovora
269 <400> SEQUENCE: 18
270 ccatgaaata cctgctgccg accgctgctg ctgggtctgct gctcctcgct gccagccgg      60
272 cgaacgct      68
275 <210> SEQ ID NO: 19
276 <211> LENGTH: 22
277 <212> TYPE: PRT
278 <213> ORGANISM: Erwinia carotovora
280 <400> SEQUENCE: 19
282 Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala
283 1      5      10      15
285 Ala Gln Pro Ala Asn Ala
286      20
288 <210> SEQ ID NO: 20
289 <211> LENGTH: 65
290 <212> TYPE: DNA
291 <213> ORGANISM: Escherichia coli
293 <400> SEQUENCE: 20
294 ccatgaaaaa gacagctatc gcgattgcag tggcactggc tggtttcgct accgtagcgc      60

```

## RAW SEQUENCE LISTING

DATE: 07/10/2001

PATENT APPLICATION: US/09/875,494

TIME: 11:51:32

Input Set : A:\Pto.amc

Output Set: N:\CRF3\07102001\I875494.raw

```

296 aggcc
299 <210> SEQ ID NO: 21
300 <211> LENGTH: 21
301 <212> TYPE: PRT
302 <213> ORGANISM: Escherichia coli
304 <400> SEQUENCE: 21
306 Met Lys Lys Thr Ala Ile Ala Ile Ala Val Ala Leu Ala Gly Phe Ala
307 1 5 10 15
309 Thr Val Ala Gln Ala
310 20
312 <210> SEQ ID NO: 22
313 <211> LENGTH: 33
314 <212> TYPE: DNA
C--> 315 <213> ORGANISM: Artificial
317 <220> FEATURE:
318 <223> OTHER INFORMATION: synthetic
320 <400> SEQUENCE: 22
321 acccggttttt tgggctaaca ggaggaatta acc
324 <210> SEQ ID NO: 23
325 <211> LENGTH: 24
326 <212> TYPE: PRT
C--> 327 <213> ORGANISM: Artificial
329 <220> FEATURE:
330 <223> OTHER INFORMATION: synthetic peptide
332 <400> SEQUENCE: 23
334 Met Ala Lys Lys Asn Ser Thr Leu Leu Val Ala Val Ala Ala Leu Ile
335 1 5 10 15
337 Phe Met Ala Gly Arg Ala Leu Ala
338 20
340 <210> SEQ ID NO: 24
341 <211> LENGTH: 72
342 <212> TYPE: DNA
C--> 343 <213> ORGANISM: Artificial
345 <220> FEATURE:
346 <223> OTHER INFORMATION: synthetic
348 <400> SEQUENCE: 24
349 atggctaataa agaactccac cctgctcggt gcagtagctg cgctgatctt catggccgga
351 agggccttgg cc
354 <210> SEQ ID NO: 25
355 <211> LENGTH: 16
356 <212> TYPE: DNA
C--> 357 <213> ORGANISM: Artificial
359 <220> FEATURE:
360 <223> OTHER INFORMATION: synthetic
362 <400> SEQUENCE: 25
363 taatgaattc aagctt

```

## VERIFICATION SUMMARY

PATENT APPLICATION: US/09/875,494

DATE: 07/10/2001

TIME: 11:51:33

Input Set : A:\Pto.amc

Output Set: N:\CRF3\07102001\I875494.raw

L:13 M:270 C: Current Application Number differs, Replaced Current Application No  
L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:23 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:1  
L:39 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:2  
L:55 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:3  
L:71 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:4  
L:87 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:5  
L:101 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:6  
L:115 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:7  
L:129 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:8  
L:143 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:9  
L:157 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:10  
L:171 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:11  
L:185 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:12  
L:199 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:13  
L:213 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:14  
L:229 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:15  
L:315 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:22  
L:327 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:23  
L:343 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:24  
L:357 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:25

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/875,494

DATE: 07/10/2001

TIME: 11:51:07

Input Set : A:\875494.txt

Output Set: N:\CRF3\07102001\I875494.raw

Does Not Comply  
Corrected Diskette Needed

5 <110> APPLICANT: Schmidt, Brian  
 6 Chen, Tseng-hui  
 8 <120> TITLE OF INVENTION: Novel Leader Peptides for Enhancing Secretion of Recombinant  
 Protein from  
 9 a Host Cell  
 11 <130> FILE REFERENCE: COUL-012-01  
 C--> 13 <140> CURRENT APPLICATION NUMBER: US/09/875,494  
 C--> 13 <141> CURRENT FILING DATE: 2001-06-05  
 13 <150> PRIOR APPLICATION NUMBER: US 60/209,517  
 14 <151> PRIOR FILING DATE: 2000-06-05  
 16 <160> NUMBER OF SEQ ID NOS: 25  
 18 <170> SOFTWARE: PatentIn version 3.0

## ERRORED SEQUENCES

354 <210> SEQ ID NO: 25  
 355 <211> LENGTH: 16  
 356 <212> TYPE: DNA  
 C--> 357 <213> ORGANISM: Artificial  
 359 <220> FEATURE:  
 360 <223> OTHER INFORMATION: synthetic  
 362 <400> SEQUENCE: 25  
 363 taatgaattc aagctt  
 E--> 366 10.  
 E--> 368 1.

16



## VERIFICATION SUMMARY

PATENT APPLICATION: US/09/875,494

DATE: 07/10/2001

TIME: 11:51:08

Input Set : A:\875494.txt

Output Set: N:\CRF3\07102001\I875494.raw

L:13 M:270 C: Current Application Number differs, Replaced Current Application No  
L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:23 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:1  
L:39 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:2  
L:55 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:3  
L:71 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:4  
L:87 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:5  
L:101 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:6  
L:115 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:7  
L:129 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:8  
L:143 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:9  
L:157 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:10  
L:171 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:11  
L:185 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:12  
L:199 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:13  
L:213 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:14  
L:229 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:15  
L:315 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:22  
L:327 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:23  
L:343 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:24  
L:357 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:25  
L:366 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:17 SEQ:25  
L:366 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:3  
M:254 Repeated in SeqNo=25  
L:368 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:2  
L:368 M:252 E: No. of Seq. differs, <211>LENGTH:Input:16 Found:18 SEQ:25